

University of Groningen

## New applications of dynamic combinatorial chemistry to medicinal chemistry

Hartman, Alwin

DOI:  
[10.33612/diss.102259269](https://doi.org/10.33612/diss.102259269)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2019

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Hartman, A. (2019). *New applications of dynamic combinatorial chemistry to medicinal chemistry*. [Thesis fully internal (DIV), University of Groningen]. Rijksuniversiteit Groningen.  
<https://doi.org/10.33612/diss.102259269>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Chapter 1

---

## Introduction to dynamic combinatorial chemistry

---

*Dynamic combinatorial chemistry (DCC) is a powerful tool to identify bioactive compounds. This efficient technique allows the target to select its own binders and circumvents the need for synthesis and biochemical evaluation of all individual derivatives. An ever-increasing number of publications report the use of DCC on biologically relevant targets. The work here complements reviews by focusing on the experimental protocol and giving detailed examples of essential steps and factors that need to be considered, such as protein stability, buffer composition and cosolvents.*

*This chapter has been published as a review article:*

A. M. Hartman, R. M. Gierse, A. K. H. Hirsch, *Eur. J. Org. Chem.* **2019**, 3581–3590.

*A. M. Hartman and R. M. Gierse contributed equally to the work in this chapter.*

*Section 1.4 was taken from the submitted review article:*

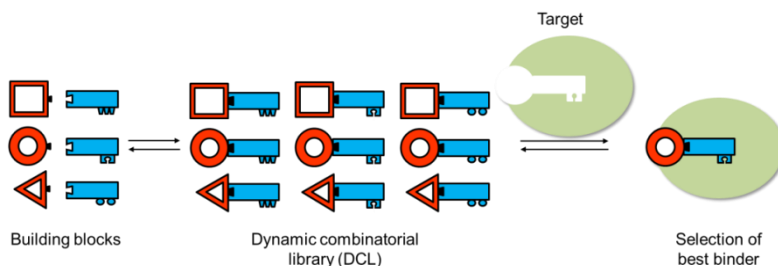
P. Kirsch, A. M. Hartman, A. K. H. Hirsch, M. Empting,

*A. M. Hartman was involved in writing the DCC example and editing the review.*

## 1.1 Introduction

Since its dawn more than two decades ago, combinatorial chemistry approaches<sup>[1–5]</sup> have developed into target-directed dynamic combinatorial chemistry (tdDCC) and have matured as a hit-identification tool. There has been an increasing number of work published in this niche of supramolecular chemistry.<sup>[6–11]</sup> A growing number of groups have shown the general applicability and scope of tdDCC for the identification of modulators of targets.<sup>[6,12–20]</sup> tdDCC refers to general pharmacologically relevant targets which next to proteins also include DNA and RNA, whereas protein-templated DCC (ptDCC) only refers to proteins. Several reviews and book chapters on tdDCC have been published in recent years.<sup>[21–23]</sup> This chapter covers our work on ptDCC and provides the key features of our protocol, explaining the essential steps in designing a successful ptDCC experiment.

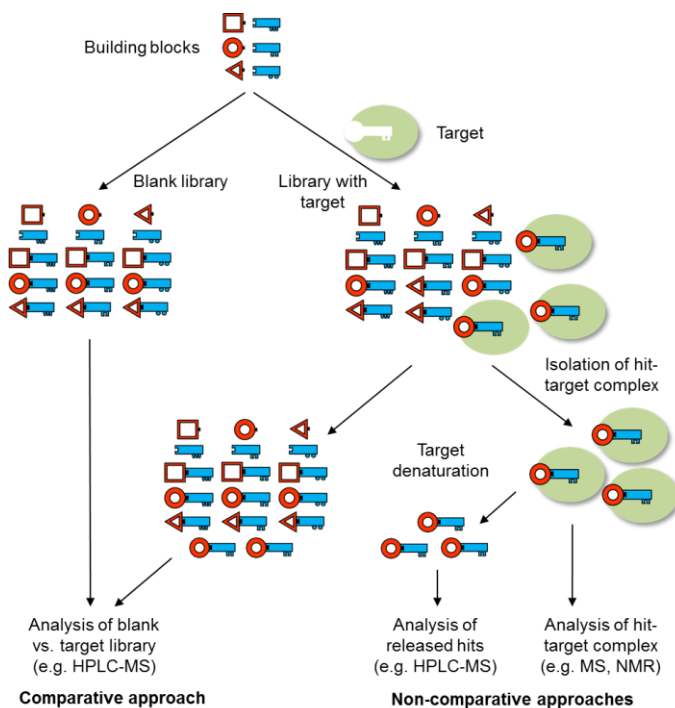
Carefully chosen building blocks are connected in a reversible manner *via* covalent or noncovalent bonds to form a dynamic combinatorial library (DCL) (Figure 1). Biocompatibility, pH dependence, temperature, solubility and stability of the components are important factors, which should be taken into account. The ideal DCLs do not require cosolvents, however, it can occur that the formed products have a lower solubility than the building blocks and in order to keep all compounds in solution, a cosolvent such as dimethyl sulfoxide (DMSO) is commonly used. Precipitation of DCL components could lead to an undesired shift in the equilibrium. By contrast, a desired shift of the equilibrium can be obtained by the addition of an external stimulus, such as a protein target. There are in general two different approaches that can be followed in ptDCC: ‘adaptive DCC’, in which the target is present during the formation of the DCL and ‘pre-equilibrated DCC’, in which the target is added after the DCL is established. An advantage of pre-equilibrated DCC is that the exchange chemistry can be applied in conditions which are not tolerated by the protein. A disadvantage is that the screening step is performed under static conditions and no amplification effects can be observed since the protein does not alter the equilibrium.



**Figure 1.** Schematic representation of target-directed dynamic combinatorial chemistry.

In ptDCC, the member(s) of DCLs, which bind best will be amplified, leading to an increase in their concentration compared to a control reaction without the external stimulus. These binders can then be further evaluated for their biochemical properties.

To enable a comparative analysis of DCLs, a blank reaction, without the target, should be run concurrent with a templated reaction. Another approach of DCC is non-comparative, in which the hits can be analysed in complex with the target or after being released from the target. There are different techniques that can be used to analyse the DCLs: liquid and size-exclusion chromatography coupled to mass spectrometry, NMR spectroscopy, fluorescence spectroscopy and X-ray crystallography. Figure 2 illustrates the comparative approach versus the non-comparative approach, which can be adopted in DCC. The reaction mixture can be ‘frozen’, in order to prevent the library from re-equilibrating during the analysis. In the case of acylhydrazone chemistry, this can be achieved by an increase in pH. Denaturation by heat, addition of a solvent or (ultra-fast) centrifugation ensures that all binders are released from the protein before analysis.



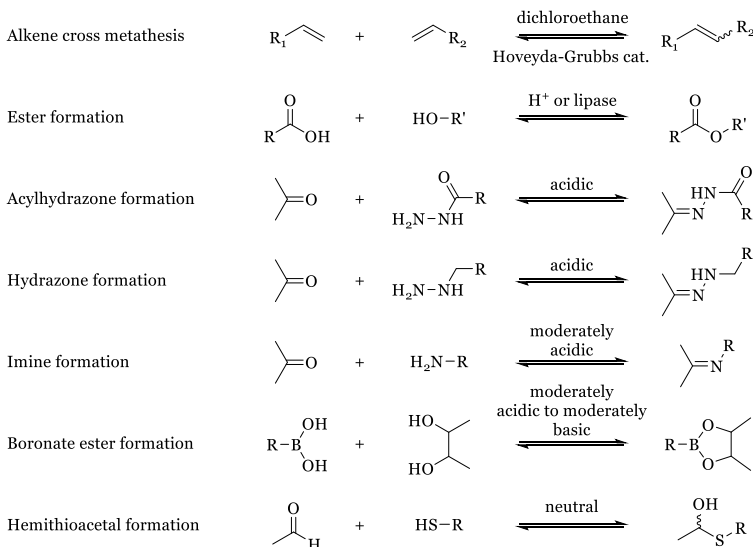
**Figure 2.** DCC approaches: comparative and non-comparative. In the comparative approach the library in presence of a target is compared to the library in absence of the target. In the non-comparative approaches, the hit–target complexes will be separated from the mixture and analysed as a complex or as released hits. The figure was adapted from Frei et al.<sup>[21]</sup>

### 1.1.1 Reversible reactions suitable for DCC

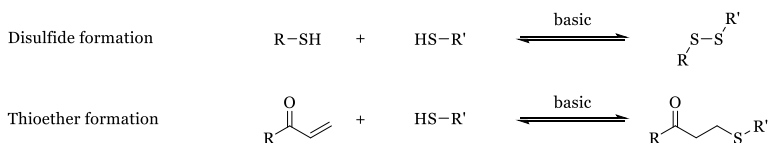
Only a limited number of reversible reactions have been used thus far, they are summarised in Scheme 1. One of the most frequently used reactions is the (acyl-)hydrazone formation, which combines ketone or aldehyde building blocks with (acyl-)hydrazides. This condensation reaction can take place in water, making it biocompatible.<sup>[24]</sup> The synthesis of the building blocks is generally straightforward or they may be commercially available.

At physiological conditions, neutral pH and room temperature, acylhydrazone formation and exchange are relatively slow. At acidic pH, the equilibrium is reached rapidly. However, Greaney and coworkers have shown that the pH dependence can be influenced by the addition of a nucleophilic catalyst. They were able to reach equilibrium reasonably fast at a comparatively high pH of 6.2 by using aniline, as a nucleophilic catalyst.<sup>[12]</sup> Previously Dawson and coworkers have shown that aniline could serve as a catalyst for acylhydrazone formation and oxime ligation.<sup>[25,26]</sup> Derivatives of aniline, which bear substituents at the aryl ring, are even more effective catalysts.<sup>[27]</sup>

The acylhydrazone linkage is reversible under acidic conditions and stable against hydrolysis at physiological pH values, allowing for the ‘freezing’ of the reversible reaction upon increasing the pH.<sup>[24]</sup>



**Scheme 1.** Reversible reactions used in target-directed DCC to identify bioactive compounds. Scheme was adapted from Van der Vlag and Hirsch.<sup>[23]</sup>



**Scheme 1 continued.** Reversible reactions used in target-directed DCC to identify bioactive compounds. Scheme was adapted from Van der Vlag and Hirsch.<sup>[23]</sup>

An overview of studies published over the past five years in the field of ptDCC is given in Table 1. It must be noted that much more work has been published applying DCC for the formation of diverse libraries in the drug-discovery process. For example the coupling of DCC to DNA-encoded libraries, creating so called DNA-encoded dynamic combinatorial chemical libraries (EDCCLs). Iminobiotin and homotetrameric streptavidin were used as a model system to identify a bidentate protein/ligand interaction. The addition of an external stimulus, for example a target protein, can shift the thermodynamic equilibrium and hence a DNA amplification can be observed after sequencing.<sup>[28]</sup>

**Table 1.** Protein-templated DCC studies reported over the past five years, in which a target was used as a template to influence the equilibrium. Therefore, only articles using an adaptive approach are listed, pre-equilibrated DCC examples are omitted.<sup>[29–31]</sup> The table is adapted from Frei et al and complemented.<sup>[21]</sup>

Target	Reversible reaction	Analysis	Library size	Equilibration time	Method applied for affinity measurement	Best affinity	Ref.
Wt Tau RNA	Disulfide	HPLC-MS and NMR	21	2 days	Fluorescence titration	$EC_{50} = 70 \text{ nM}$	[32]
HIV FSS RNA	Disulfide	MS	12	4 days	n.a.	n.a.	[33]
Vascular endothelial growth factor receptor (VEGFR) 2	Imine	HRMS	297	24 h	In vitro activity against cancer cell lines	$IC_{50} = 2.4 \text{ }\mu\text{M}$	[34]
Endothiapepsin	Acylhydrazide	HPLC-MS	90	20 h	Inhibition assay	$IC_{50} = 54.5 \text{ nM}$ $K_i = 25.4 \text{ nM}$	[35]
FimH	Acylhydrazide	HPLC	8	3 days	SPR	$K_D = 273 \text{ nM}$	[36]
UDP-galactose-4-epimerase	Acylhydrazide	HPLC	11	24 h	Fluorescence-based assay and MIC	$K_D = 3 \text{ }\mu\text{M}$ $MIC = 26 \text{ }\mu\text{g mL}^{-1}$	[37]

Myeloperoxidase (MPO)	Hydrazone	Activity assay	6	n.a.	in vivo activity assay	$IC_{50} = 79 \text{ nM}$	[38]
ecFabH	Acylhydrazide	$^{19}\text{F}$ -NMR	5	12 h	Enzymatic assay	$IC_{50} = 3 \text{ mM}$	[39]
Multi-protein strategy on AlkB oxygenases: FTO, ALKBH3 and ALKBH5	Acylhydrazide	DSF and HPLC	10	5 h	HPLC-based demethylase and DSF assays	$IC_{50} = 2.6 \text{ }\mu\text{M}$	[40]
<i>Trypanosoma cruzi</i> bromodomain-containing (TcBDF3)	Acylhydrazide	HPLC-MS	30	n.a.	DSF	$IC_{50} = 13\text{--}23 \text{ }\mu\text{M}$	[41]

DSF = differential scanning fluorimetry, HPLC = high-performance liquid chromatography,  $IC_{50}$  = half maximal inhibitory concentration, ITC = isothermal titration calorimetry,  $K_D$  = dissociation constant,  $K_i$  = inhibition constant, MIC = minimum inhibitory concentration, MS = mass spectrometry, n.a. = not available, NMR = nuclear magnetic resonance, SPR = surface plasmon resonance,  $T_m$  = thermal shift.

## 1.2 A closer look on the templating protein

To obtain meaningful results from DCC experiments, the quality of the input template is critical. As the equilibrium of the library shifts by the templating effect of the added protein sample, it should consist of the target protein as close to its native state as possible. The quantity of the used template depends on the protein target, there are reported successful DCC projects with 0.1 to 1.5 equivalents of protein. [29,42] DCC experiments are also possible with a mixture of proteins, but a well-defined sample eases up downstream data analysis and reduces the number of false positives for the desired target. [40] The condition of the protein sample depends on various variables. For DCC experiments the purity, concentration, tertiary and quaternary structure of the protein, additives and contaminations, as well as the pH-value are of particular importance. During the experiment, which can take up to several days, protein degradation and precipitation could occur. The tests described herein should give an overview and help to choose suitable experimental conditions to plan new DCC experiments. In the next paragraphs, we will briefly discuss the influence of those factors and suitable analytical methods to monitor them.

### 1.2.1 Purity

In the case of a mixed or impure protein sample, there might be several templated reactions proceeding in parallel. It is impossible to differentiate between a small fraction of the sample showing a strong template effect and a large fraction of the

protein pool showing only a weak amplification of a binder. This will result in overlapping data, which are difficult to analyse, and may result in false positives. We therefore recommend starting with the highest protein purity available.

### **1.2.2 Stability**

Not only the initial state, but also the stability of the templating protein during the reaction should be checked by preliminary tests before conducting a DCC experiment. The time span over which a DCC experiment, pre-equilibrated or adaptive, is monitored can vary. It depends on the reaction rate and concentration and should ideally be monitored until the library reaches an equilibrium state. Usually, the DCL reaches a new equilibrium within the first few days, depending on the reversible reaction and conditions (Table 1). However, if the protein is stable for longer periods of time, longer equilibration times are possible, for example up to 20 days for the very stable protease endothiapepsin (see Section 1.2.4).<sup>[24]</sup>

It is important that the protein is not precipitating or degrading during the experiment. Precipitation of the protein will remove the template from the solution. Denaturation of the template will lead to entirely new templates, which would affect the equilibrium state of the DCL. This can lead to random and irreproducible amplification of compounds by the unordered protein and a decrease of initially already amplified best binders of the native template. If the protein target is labile, it is therefore necessary to follow the reaction over time to identify the temporary, templated equilibrium of the DCC library. In this, compounds amplified by the native state of the template can be found.

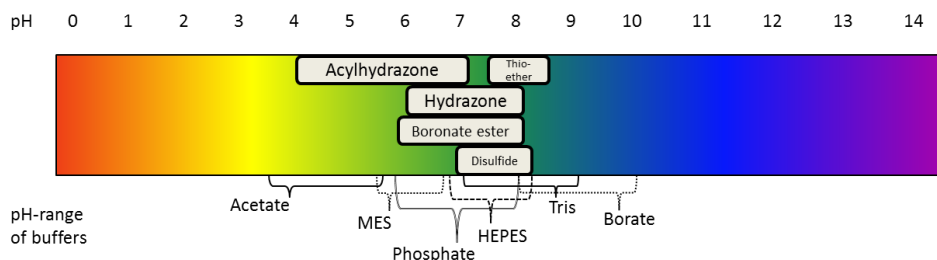
Eventually, after prolonged incubation time, nearly every protein will degrade and, by this, change the equilibrium of the DCL again. Compounds amplified in this step can be ignored, as they were not templated by the native protein. Observation of the DCC experiment for longer timeframes than the template integrity can be guaranteed is therefore of no use.

### **1.2.3 Buffer and pH**

When choosing a buffer for DCC experiments, several different requirements have to be met. Attention should be paid to possible side reactions with the DCL or chelation effects. For example, Tris-buffer could form imines with aldehyde building blocks, which might influence the formation of the DCL. Some stabilization of the protein is beneficial, but strong interactions of the buffer with the target protein should also be avoided, for instance, a phosphate buffer for a phosphate binding protein. The phosphate could compete with possible binders; possible effects of competition are discussed in more detail in Section 1.2.5. So far, in most cases common buffer systems have been used, which are shown in



Table 2 and Scheme 2. The choice of buffer is, however, not limited to the established systems.



**Scheme 2.** Example of possible buffers and the pH ranges of reactions used in DCC experiments.

**Table 2.** Buffers commonly used in different DCC reactions. \*Tris buffer requires special attention.

Reaction	Buffer described in literature
Acylhydrazone formation <sup>[6,24,36]</sup>	Ammonium- and Sodium acetate, Phosphate, Tris*
Hydrazone formation <sup>[43,44]</sup>	Phosphate, Tris*
Disulfide <sup>[45,46]</sup>	Phosphate, Borate
Thioether <sup>[47]</sup>	Water/DMSO
Imine <sup>[13]</sup>	Water
Boronate ester <sup>[14,48]</sup>	Ammonium acetate, Water

For many protein targets, the stability at room temperature and the optimal buffer conditions are not known. We therefore recommend determining these conditions prior to performing DCC experiments. As several interdependent factors, like pH, buffer, ionic strength and ions influence the stability of a protein, it is difficult to suggest a stepwise flow scheme for the determination of the ideal buffer composition for a given protein.<sup>[28]</sup> Not only the protein but also the exchange chemistry might be affected significantly by varying these parameters. We propose to first measure the effect of pH, buffer and ionic strength over a wide

range in parallel. Afterwards, a small selection (2 to 5) of the most stabilizing combinations can be evaluated for their long-term effect on the protein. Subsequently, the best condition will then be used to determine the influence of DMSO (Section 2.6) and, if of interest, additives (Section 2.5). The selection of the initial buffers could be broadened, in case no suitable condition was found.

Two or more buffers should be screened per pH value to distinguish the influence of the buffer component and the pH value on the stability of the protein. It is also possible to use a so-called “superbuffer”,<sup>[49]</sup> a mix of three or more buffer components, enabling the adjustment of a wide pH-range, without changing the buffer composition or concentration.

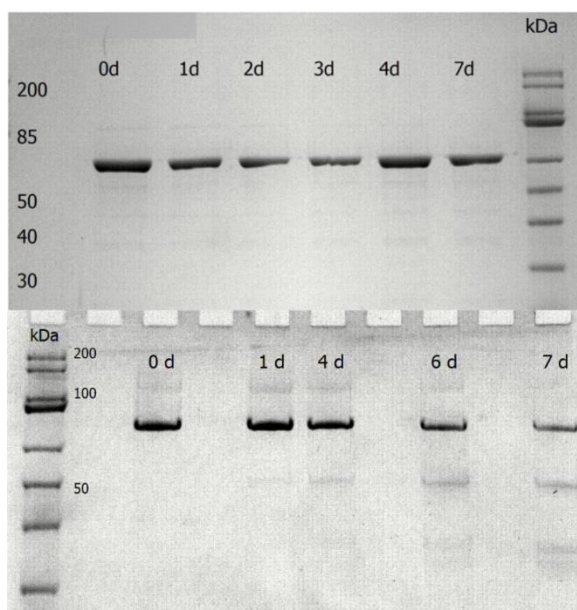
The effect of the buffer components on a protein can be measured in a straightforward way, by determining the melting point of the target protein *via* a thermal-shift assay / differential scanning fluorimetry (TSA /DSF).<sup>[28]</sup> In this method, the protein is incubated together with a lipophilic dye, for example sypro orange. The dye shows an increase in fluorescence after binding to the hydrophobic parts of a protein. These are often located at the inside of a protein and become exposed during temperature-induced unfolding/melting. The temperature-dependent increase in fluorescence can be measured in a RT-PCR apparatus and yields the  $T_m$  of the protein.

Other methods, like DSC, ITC and CD (differential scanning calorimetry, isothermal titration calorimetry and circular dichroism spectroscopy) and the determination of melting points by CD could also be used to gain information on the interaction and possible stabilization of the protein with its buffer, but require a high amount of protein and/or long measurement time. The TSA, however, offers high throughput and a short assay time, together with already several published or commercially available kits in 96-well format.<sup>[50,51]</sup> These kits were originally intended to screen for optimal crystallization conditions and cover several stability-influencing conditions. When performing DCC experiments, the design of an individual 96-well plate layout, tailored to the buffers and conditions compatible with the planned DCC reaction, might be useful. This is a short time investment, which might pay off quickly in the future, if ptDCC is used on several different targets.

After a DCC-compatible, stabilizing buffer condition has been identified, the protein should be checked for its long-term stability. To check for cleavage of the protein backbone an analysis by SDS-PAGE is of sufficient sensitivity (Figure 3). To determine if the protein folding is affected, TSA is again the method of choice, since the signal directly depends on the unfolding process of the protein. With prolonged degradation, the melting point decreases slightly. As a secondary effect of the degradation, the fluorescence curve can show bi- and multi-phasic melting

points and an overall decrease in signal intensity and resolution. A fully denatured enzyme will just show a decreasing fluorescence signal with no peak from protein unfolding. As controls, a fresh and a heat-treated sample of the target protein should be included in the experiment.

The tendency of a protein to precipitate is concentration-dependent. Because of this, the assays determining the protein stability should be performed with the same protein concentration that is intended to be used in the DCC experiment. If this is not possible, due to limited protein availability, the first experiments might be done with less protein. However, at least for the chosen final condition, the stability assessment should be repeated with the protein concentration that will be used in the DCC experiments.

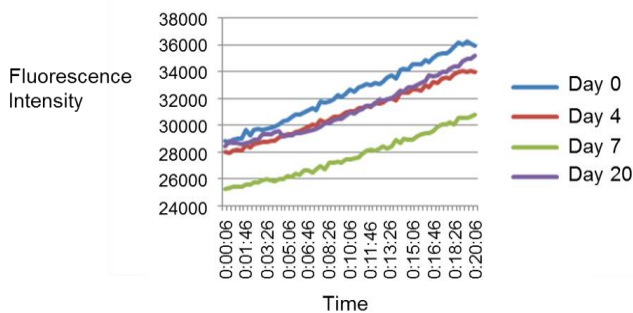


**Figure 3.** 12% SDS-PAGE of different homologues of the enzyme 5-Deoxyxylulose 5-Phosphate Synthase (DXS) after incubation at RT. The protein on the upper gel shows no sign of degradation. The second protein, shown on the lower gel, shows signs of degradation, starting already at day one with a very faint band around 50 kDa. From day 6 on a decrease of the main protein band also becomes clear. In the top left corner a gel-label was removed using image processing software.

### 1.2.4 Functional enzyme assay

For enzymatic protein targets, a functional assay can be used instead of TSA and PAGE measurements for the assessment of long-term stability. The analysis of activity data of a functional assay to determine the best experimental conditions of the DCC experiments leaves less room for interpretation than the analysis of the results of a melting-point analysis. Therefore, if a functional assay is available, and the enzyme is showing catalytic activity in the desired pH range of the DCC reaction, the activity assay should be the method of choice.

In a previous study from 2014, we could monitor the activity of the target protein endothiapepsin by performing a fluorescence-based assay (Figure 4). The pH-optimum of endothiapepsin is 4.5, and the enzymatic activity was not affected even after 20 days incubation at RT and a pH of 4.6. Considering this high stability, no buffer optimization was needed.<sup>[24]</sup>



**Figure 4.** Activity of endothiapepsin, a pepsin-like aspartic protease, in a fluorescence-based assay at different time intervals of incubation at room temperature. Figure was adapted from Mondal et al.<sup>[24]</sup>

### 1.2.5 Additives and contaminations

During the purification, the protein might be in contact with different buffers and conditions. Some of the buffer components might remain bound to the protein, even after a buffer exchange. These contaminants might influence the experiment. It is therefore recommended to critically evaluate the composition of the protein sample. Not only should the protein storage buffer be evaluated, but also the origin of the sample.

Common substances that could be found in protein samples are for example imidazole as a leftover from an IMAC (immobilised metal affinity chromatography) purification step. Protein samples are often supplemented with reducing agents like 2-ME, DTT or TCEP (2-Mercaptoethanol, Dithiothreitol or Tris(2-carboxyethyl)phosphine) in concentrations up to 10 mM to keep the

protein in a reducing environment. If disulfide formation is the reversible reaction of choice, the final reducing agent concentration should be evaluated to make sure that the formation of disulfide bonds is not inhibited.

The effect of additives and contaminations is related to the volume of the protein sample used in the individual DCC experiments. This not only determines the final concentration of protein, but also the concentrations of the contaminants. If the batch-to-batch concentration of the protein varies and its volume is adjusted to reach the same final concentration in the DCC experiment, it should be noted that the concentrations and effects of the additives in the DCC experiment might vary.

Compounds that remain in the protein sample can have an influence on the DCC reaction or on the target protein. One should critically check every buffer component on possible interference with the planned exchange chemistry. Screening literature for known reactions between the DCL members and sample components can be considered. Performing a control experiment with all buffer components, in the absence of protein, can assure that no side reactions are taking place. If the exact composition of the protein sample is unknown, a small volume of the buffer might be gained by concentration of the protein using an ultrafiltration device and using the flow-through for the control experiment.

Some agents used during protein purification, such as cryoprotectants like glycerol or detergents like Tween, will interact in a non-specific way with the protein surface. From our experience, if there is no hint that they might affect the experiment, leftover cryoprotectants and detergents can be tolerated. Special care should be taken if cofactors, coenzymes or ions are supplemented during the purification process to stabilise the enzyme. The same holds true for buffer components structurally related to those supplements. Everything that binds to the targeted binding pocket is competing with the DCC library. If a natural, tight binding cofactor is present during the experiment, it could prevent the building blocks from binding and therefore also inhibit their amplification. However, the use of tight binders can be beneficial in control experiments. If a compound with a known binding site is inhibiting the formation of some previously observed binders this can be taken as a hint that the templated binders are targeting the same protein pocket.

### **1.2.6 DMSO**

Addition of a small percentage of DMSO to the reaction solution is a common practice in the design of enzymatic assays to improve the solubility of hydrophobic compounds. For biochemical assays, DMSO concentrations up to 10% are regularly used.<sup>[52]</sup>

In DCC experiments, the building blocks of the library are typically dissolved in DMSO stock solutions to enable the easy assembly of a library. Depending on the library composition and number of compounds used, the final DMSO concentration would vary. To keep the reaction conditions comparable, we recommend adding DMSO up to a concentration that can be kept constant for all experiments of a project. This fixed concentration should be evaluated and chosen beforehand, to ensure the protein tolerates it.

DMSO has a very broad range of effects on proteins, it can even decrease the solubility and induce precipitation.<sup>[53]</sup> Both, rate acceleration, as well as inhibition of the enzyme-catalysed reaction by DMSO have been observed. An influence of already low percentages of DMSO on the enzymatic activity often hints to DMSO acting as an unspecific effector, interacting with the active site of the enzyme.<sup>[54]</sup> If the enzymatic activity is reduced by DMSO at higher concentrations (>10% DMSO), it is often by influencing the overall protein conformation by displacing water molecules bound to the surface and unfolding the protein.<sup>[55]</sup> On the other hand, there are DMSO-tolerant enzymes known which show activity up to 80% DMSO.<sup>[54]</sup> Enzyme activity assays are the method of choice to estimate the effect of DMSO on an enzyme. If no activity assay is available, the effect of DMSO could also be measured using TSA, however, interactions with the active site are difficult to detect with this method. We often observe a small effect on the  $T_m$  of a protein, but a strong effect on the enzymatic activity. Taken together, the DMSO concentration has several effects on the protein structure. The benefits of DMSO addition need to be weighed against the risk of creating an artificial enzymatic fold, which could amplify compounds that would not bind under native conditions. Therefore, the DMSO concentration should be as low as possible, in our lab up to 5% are regularly used.

### **1.2.7 Temperature**

To speed up the rate at which the DCL reaches equilibrium, the experiments are normally performed at room temperature. For labile proteins, a lower reaction temperature may be necessary, which can improve the stability of the proteins. At the same time, the equilibration rate is decreased, leading to a prolonged incubation time. The optimal temperature for protein stability in DCC could vary from enzyme to enzyme and thus needs to be evaluated in each individual case but room temperature is used in most cases.

## **1.3 Setting up a ptDCC experiment**

When crystal structures are available, or even cocrystal structures, a structure-based approach can be undertaken to design promising building blocks. In this case also non-binders could be designed as control elements, which are not supposed to emerge as hits. The type of reversible linkage should be carefully

selected because it influences the molecular recognition by the target. For example, the acylhydrazone linkage resembles the amide functionality; and features hydrogen-bond donors and acceptors. We showed that by combining DCC with de novo structure-based design, the risks associated with this attractive approach are reduced.<sup>[24]</sup>

### 1.3.1 Formation of the DCLs

The building blocks might have to be dissolved in DMSO, allowing them as well as the formed products to stay soluble in the final mixture. In principle, they could also be dissolved in the desired buffer, which would be most ideal. In 2014, we coupled DCC to saturation-transfer difference (STD)-NMR spectroscopy, which requires lower concentrations of protein than a general DCC experiment (Table 3). STD-NMR spectroscopy enables selection of the binders from the DCL, since the intensity of these signals is stronger due to a more efficient saturation transfer. As a result of only observing binders, STD-NMR spectra cannot be used to determine concentrations of DCL members and therefore amplification cannot be calculated. In follow-up experiments, it is possible to determine the  $K_D$  value of a ligand *via* STD-NMR or other biophysical assays.<sup>[56]</sup>

The ratio of hydrazides versus aldehydes should allow for the formation of all possible products, therefore at least one equivalent of each hydrazide per aldehyde should be used. For example, if three aldehydes are used then at least three equivalents of each hydrazide should be added, making sure that there is an excess of hydrazides. When required, a nucleophilic catalyst like aniline could be added. The most frequently used concentration of DMSO lies around 5–10%.

**Table 3.** General protocol for DCC and protocol for DCC coupled to  $^1\text{H}$ -STD-NMR. \* Aniline or another nucleophilic catalyst could be added when required. \*\* In a control experiment, no protein is added. \*\*\* Buffer conditions to guarantee protein stability should be determined a priori.

	Final concentration in general DCC	Final concentration used in DCC coupled to $^1\text{H}$ -STD-NMR <sup>[24]</sup>
Aldehyde	0.1 mM	0.4 mM
Hydrazide	0.1–0.3 mM	1 mM (for each of the five hydrazides)
DMSO	5–10 %	5–10 %
Aniline*	10 mM	–
Protein**	10–100 $\mu\text{M}$	4 $\mu\text{M}$
Buffer***	0.1 M	Ammonium acetate in $\text{D}_2\text{O}$ (0.1 M, pH 4.6)
pH*	Acidic–neutral	pH 4.6

Control experiments should be considered, which should clarify where binding of molecules to the protein occurs and if it is specific or unspecific. This could for example be performed by the addition of a known inhibitor. If the previously observed amplification is not observed any longer, then the hit compounds are competitive binders. Based on the work of Danieli et al., B. Ernst and coworkers propose that the use of bovine serum albumin (BSA), as a negative control template for which no amplification is expected since the binding pocket is different, is not a good control since it could influence the library composition, whilst the use of a competitive inhibitor is better. BSA has been used in DCC to show that the applied library only gives hits with the real target and that BSA would yield the same result as the blank.<sup>[21,57]</sup> BSA is commonly known for its stability and was thought not to interfere with biological reactions, however recently DCC experiments have even been used to target BSA.<sup>[58]</sup>

### 1.3.2 Analysis of the DCLs

Different techniques such as fluorescence-polarization, SPR, ITC, MST, STD-NMR, crystallography and others can be used to evaluate and possibly optimise obtained hits. We and Rademann and coworkers have reviewed the analytical methods used in protein-templated dynamic combinatorial chemistry to detect hit compounds.<sup>[23][59]</sup>

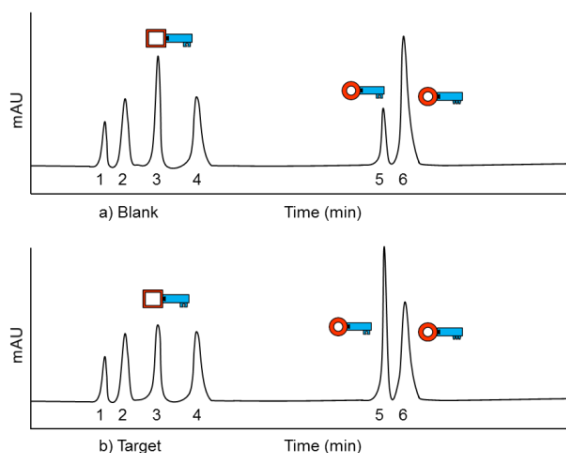
A commonly applied method to analyse DCC experiments is the recording of HPLC-MS chromatograms of the libraries. As an illustrative example of the comparative approach, we drew HPLC chromatograms of a blank library and a target library (Figure 5). When we compare both chromatograms, we see that peak number five has increased in the library containing the target, whereas peaks three and six have decreased. The total amount of building blocks stays the same, only the equilibrium can be shifted towards one or more products.

In order to accurately determine the amplification or decrease of peaks, their relative peak areas (RPA) should be compared. The fictional RPAs of both chromatograms in Figure 5 are given in Table 4. The amplification factor in percentage can be calculated by equation 1, where the amplification factor in ‘fold’ is given by equation 2. Using these two equations, the product at peak five has increased by 100% or twofold. Frei et al. report on a particularly thorough analysis of a DCL using the lectin FimH as a target, using HPLC analysis with an optimised DCC protocol.<sup>[36]</sup>

$$\text{Equation 1: amplification factor (\%)} = \frac{RPA_{\text{target}} - RPA_{\text{blank}}}{RPA_{\text{blank}}} * 100\%$$

$$\text{Equation 2: amplification fold} = \frac{\text{New}}{\text{Old}}$$





**Figure 5.** Schematic example of HPLC chromatograms: a) blank library chromatogram, b) target library chromatogram.

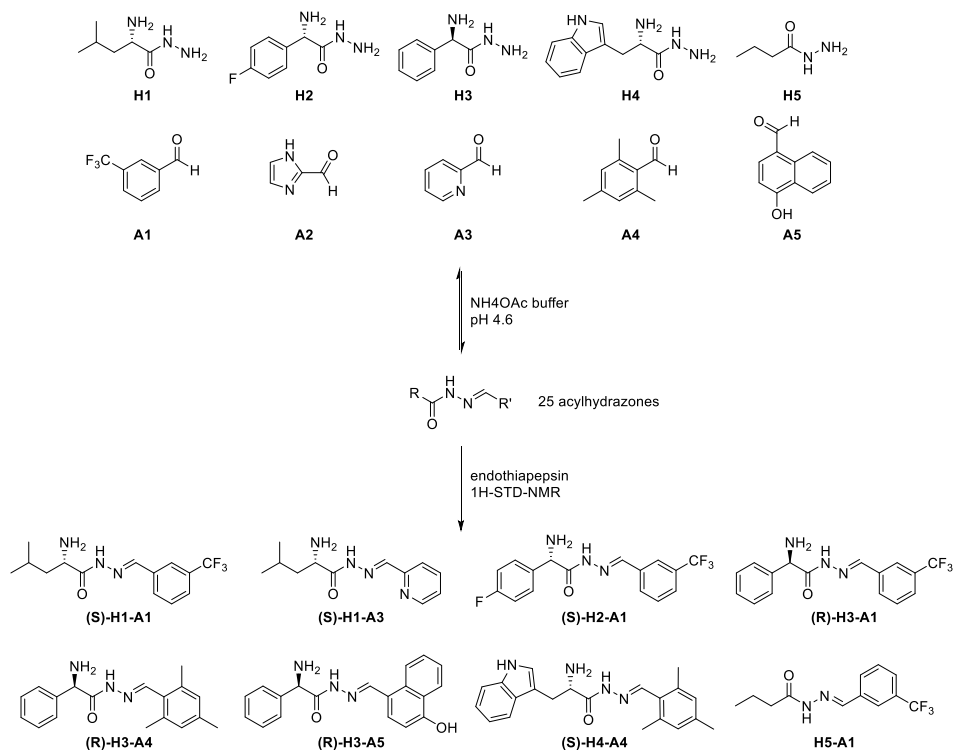
**Table 4.** Example of relative peak areas (RPA) obtained from HPLC chromatograms from Figure 4.

Peak number	Relative peak area in blank (%)	Relative peak area in target (%)	Amplification in %	Amplification in 'fold'
1	10	10	-	1
2	15	15	-	1
3	20	16	-20%	0.8
4	16	16	-	1
5	12	24	100%	2
6	27	19	-30%	0.7
Total	100%	100%		

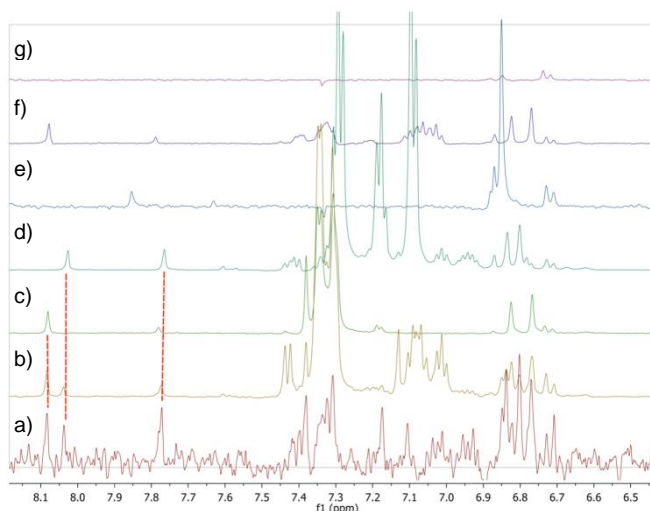
### 1.3.3 DCL analysed with STD-NMR spectroscopy

Inspired by the work of Ramström and coworkers<sup>[20]</sup>, we analysed the formed DCLs by STD-NMR spectroscopy (Scheme 3). We used the model enzyme endothiapepsin as target. As a control with a known binder we used saquinavir ( $K_i = 48$  nM), a potent peptidic inhibitor, to differentiate specific from nonspecific binding. Each sub-library contained all five hydrazides and one of the aldehyde building blocks and was allowed to equilibrate for 24 hours before adding the target. By analysing the imine-type proton signals of the acylhydrazone products in the  $^1\text{H}$ -STD-NMR spectra (Figure 6) we identified in total eight binders. To confirm the results from STD-NMR, we performed an enzyme-inhibition assay and showed that the hits were inhibitors with  $IC_{50}$  values ranging from 12.8  $\mu\text{M}$  to 365  $\mu\text{M}$ . The high hit rate in this publication is due to use of five sublibraries detecting the best binder of each library, whereas in a regular ptDCC setup only the overall best binders will be discovered. In addition, the high hit rate is also a

result of the synergistic combination of de novo structure based drug design (SBDD) and DCC. In STD-NMR the protein is used as a tool to analyse the library, whereas in a ptDCC experiment the protein influences the equilibrium and hence the concentrations.



**Scheme 3.** Formation of dynamic combinatorial library and enzymatic selection of the best binders by 1H-STD-NMR analysis; Scheme adapted from Mondal et al.<sup>[24]</sup>



**Figure 6.** DCL generated from **H1–5 + A4**: (aromatic region) a)  $^1\text{H}$ -STD-NMR spectrum of **H1–5 + A4**, b)  $^1\text{H}$ -NMR spectrum of **H1–5 + A4**, c)  $^1\text{H}$ -NMR spectrum of **H3+A4**, d)  $^1\text{H}$ -NMR spectrum of **H4+A4** (2 singlets correspond to the *E/Z* isomers), e) **H1+A4**, f) **H2+H4** and g) **H5+A4**. Figure was adapted from Mondal et al.<sup>[24]</sup>

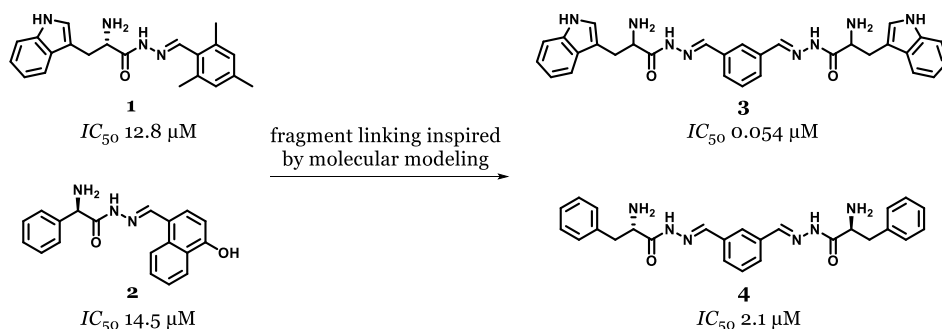
### 1.3.4 How to proceed after obtaining hits

Having obtained a validated hit, identified by de novo structure-based drug design in combination with DCC and STD-NMR, we have used a structure-based design approach to improve the molecular recognition by the target.<sup>[60]</sup> In this specific case, we were fortunate to have an x-ray crystal structure of the target endothiapepsin in complex with the hit. If this is not the case, optimization is still possible, relying on structure–activity relationships.

## 1.4 DCC in a synergistic combination with fragment linking

For fragment linking, two or more fragments have to bind to different but adjacent sites of the enzyme active site.<sup>[61]</sup> This approach introduces one additional component into the ligand system: a linker moiety. Finding the right linker motif, whilst maintaining the binding poses of both fragments, which orients the individual fragment units in the favourable geometry in relation to each other without introducing too much flexibility, can be challenging. The combination of two fragments with rather low affinity could result in significantly higher affinity and has the potential to result in ‘superadditive’ contributions of both binding motifs. The challenge in fragment linking is the exploration of the binding mode of both fragments and the identification of an optimal linker fitting in between. Only in this case, the overall reduced so-called ‘rigid body entropy’

translates into synergistically improved affinity. By binding of a fragment to a target protein, rotational and translational entropy is lost. This entropy penalty has to be overcompensated by attractive interactions formed between the ligand and the target. When two fragments bind in parallel to adjacent sites, each has to pay this entropy penalty. When these two fragments are linked together in an ideal way, the resulting singular compound only encounters the loss of rigid body entropy once. Hence, the observed affinity will be much greater than only the sum of the individual affinities.<sup>[62]</sup> The additional binding energy gained, is often also referred to as linker energy. To overcome the challenges associated with fragment linking, we pioneered a synergistic combination with DCC. For this proof-of-concept study, we again used the model enzyme endothiapepsin.<sup>[35]</sup> X-ray crystal structures of endothiapepsin in complex with fragment inhibitors **1** and **2** (PDB IDs: 4KUP and 3T7P) identified by DCC were used as a starting point for fragment linking studies facilitated by DCC. Hits **1** and **2** display  $IC_{50}$  values of 12.8  $\mu$ M and 14.5  $\mu$ M and LEs of 0.27 and 0.29, respectively. The linking of **1** and **2** should generate an inhibitor that occupies two binding pockets of endothiapepsin (Figure 7).



**Figure 7.** Structures of hits **1** and **2** and linked bisacylhydrazone linked inhibitors **3** and **4**.<sup>[35]</sup>

The homo-bis-acylhydrazones **3** and **4** were hits from the DCC experiments and were synthesised and evaluated accordingly. Compared to compound **2**, the potency of inhibitor **3** was increased 240-fold, yielding an  $IC_{50}$  value of 0.054  $\mu$ M and a LE value of 0.29. For inhibitor **4** an  $IC_{50}$  of 2.1  $\mu$ M and a LE value of 0.25 was determined (Figure 7).<sup>[35]</sup> Obviously, only the symmetric linking modality resulted in efficient cooperative binding.

## 1.5 Conclusions

There are a number of steps which should be carefully taken into account, in order to obtain active hits by DCC. If information on the target is available, e.g. a crystal-structure, one could consider a structure-based design when choosing the building blocks. The type of reversible linkage to be used can be chosen at this stage. Conditions necessary for the equilibration to take place should be compatible with the target. After establishing conditions, which will ensure the target remains folded, the actual DCC experiment can be started. To do so, stock solutions of building blocks, catalyst and protein should be prepared. The formed DCLs can be analysed by different techniques such as STD-NMR or HPLC-MS. Compounds that have been selected by the target, and their biochemical properties should be evaluated and possibly optimised in further studies.

## 1.6 Outline of this thesis

Dynamic combinatorial chemistry (DCC) has evolved over the past decades from a tool to easily generate a pool of derivatives to an efficient technique to find hit compounds when applied on a target. To be able to use target-directed DCC (tdDCC), a number of criteria must be met ranging from biocompatibility, solubility, stability, pH dependence to temperature and type of reversible reaction.

The protein family of 14-3-3 has been selected as a target, since it allows for all of these criteria to be met. 14-3-3 proteins are involved in protein-protein interactions (PPIs) in many different biological processes, ranging from diseases to cell-cycle control and signal transduction. Modulating 14-3-3 proteins for binding is therefore an important class of research.

The first aim of this thesis is applying DCC on biological relevant targets; such as 14-3-3 proteins and glucansucrases. The second aim of this thesis is on extending the list of reversible reactions, which can be applied in tdDCC. This allows medicinal chemists more freedom, by being able to use different scaffolds. Therefore, the overall objective is to find new applications of DCC to medicinal chemistry.

In chapter 3, we apply tdDCC on the 14-3-3 $\zeta$  isoform. We use the PPI complex of 14-3-3 $\zeta$ /synaptopodin in acylhydrazone-based DCC, aiming to find small drug-like molecules which can stabilise this PPI.

Chapter 4 describes the application of tdDCC on a glucantransferase, which is found to be causative for adhesion of bacteria to the tooth enamel, which can lead to dental caries.

Finally, in chapters 5 and 6 the design and first applications of new scaffolds for tdDCC is described. In chapter 5, the application of nitron-based DCC with endothiapepsin is evaluated. And in chapter 6, the thiazolidine scaffold is investigated and DCC conditions optimised for endothiapepsin.

## 1.7 References

- [1] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, *354*, 82–84.
- [2] R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley, J. H. Cuervo, *Nature* **1991**, *354*, 84–86.
- [3] R. A. Houghten, *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 5131–5135.
- [4] H. M. Geysen, R. H. Meloen, S. J. Barteling, *Proc. Natl. Acad. Sci.* **1984**, *81*, 3998 LP – 4002.
- [5] R. Frank, W. Heikens, G. Heisterberg-Moutsis, H. Blöcker, *Nucl. Acid. Res.* **1983**, *11*, 4365–4377.
- [6] I. Huc, J.-M. Lehn, *Proc. Natl. Acad. Sci.* **1997**, *94*, 2106–2110.
- [7] M. H. Ohlmeyer, R. N. Swanson, L. W. Dillard, J. C. Reader, G. Asouline, R. Kobayashi, M. Wigler, W. C. Still, *Proc. Natl. Acad. Sci.* **1993**, *90*, 10922 LP – 10926.
- [8] S. J. Rowan, P. S. Lukeman, D. J. Reynolds, J. K. M. Sanders, *New J. Chem.* **1998**, *22*, 1015–1018.
- [9] V. A. Polyakov, M. I. Nelen, N. Nazarpak-Kandlousy, A. D. Ryabov, A. V. Eliseev, *J. Phys. Org. Chem.* **1999**, *12*, 357–363.
- [10] A. Ganesan, *Angew. Chemie Int. Ed.* **1998**, *37*, 2828–2831.
- [11] C. Karan, B. L. Miller, *Drug Discov. Today* **2000**, *5*, 67–75.
- [12] V. T. Bhat, A. M. Caniard, T. Luksch, R. Brenk, D. J. Campopiano, M. F. Greaney, *Nat. Chem.* **2010**, *2*, 490–497.
- [13] Z. Fang, W. He, X. Li, Z. Li, B. Chen, P. Ouyang, K. Guo, *Bioorganic Med. Chem. Lett.* **2013**, *23*, 5174–5177.
- [14] M. Demetriades, I. K. H. Leung, R. Chowdhury, M. C. Chan, M. A. McDonough, K. K. Yeoh, Y.-M. Tian, T. D. W. Claridge, P. J. Ratcliffe, E. C. Y. Woon, et al., *Angew. Chemie Int. Ed.* **2012**, *51*, 6672–6675.
- [15] E. C. Y. Woon, M. Demetriades, E. A. L. Bagg, W. Aik, S. M. Krylova, J. H. Y. Ma, M. Chan, L. J. Walport, D. W. Wegman, K. N. Dack, et al., *J. Med. Chem.* **2012**, *55*, 2173–2184.

- [16] S. Sakai, Y. Shigemasa, T. Sasaki, *Tetrahedron Lett.* **1997**, 38, 8145–8148.
- [17] R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Angew. Chemie Int. Ed.* **2002**, 41, 3405–3407.
- [18] R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving, S. A. Brown, *Tetrahedron* **2004**, 60, 771–780.
- [19] B. Shi, R. Stevenson, D. J. Campopiano, M. F. Greaney, *J. Am. Chem. Soc.* **2006**, 128, 8459–8467.
- [20] R. Caraballo, H. Dong, J. P. Ribeiro, J. Jiménez-Barbero, O. Ramström, *Angew. Chemie Int. Ed.* **2010**, 49, 589–593.
- [21] P. Frei, R. Hevey, B. Ernst, *Chem. - A Eur. J.* **2019**, 25, 60–73.
- [22] M. Mondal, A. K. H. Hirsch, *Chem. Soc. Rev.* **2015**, 44, 2455–2488.
- [23] R. Van der Vlag, A. K. H. Hirsch, in *Compr. Supramol. Chem.* 2, Elsevier, **2017**, pp. 487–509.
- [24] M. Mondal, N. Radeva, H. Köster, A. Park, C. Potamitis, M. Zervou, G. Klebe, A. K. H. Hirsch, *Angew. Chem. Int. Ed.* **2014**, 53, 3259–3263.
- [25] A. Dirksen, S. Dirksen, T. M. Hackeng, P. E. Dawson, *J. Am. Chem. Soc.* **2006**, 128, 15602–15603.
- [26] A. Dirksen, T. M. Hackeng, P. E. Dawson, *Angew. Chem. Int. Ed.* **2006**, 45, 7581–7584.
- [27] P. Crisalli, E. T. Kool, *J. Org. Chem.* **2013**, 78, 1184–1189.
- [28] F. V Reddavid, W. Lin, S. Lehnert, Y. Zhang, *Angew. Chemie Int. Ed.* **2015**, 54, 7924–7928.
- [29] L. Monjas, L. J. Y. M. Swier, I. Setyawati, D. J. Slotboom, A. K. H. Hirsch, *ChemMedChem* **2017**, 12, 1693–1696.
- [30] F. T. Kern, K. T. Wanner, *ChemMedChem* **2015**, 10, 396–410.
- [31] F. Kern, K. T. Wanner, *Bioorg. Med. Chem.* **2019**, 27, 1232–1245.
- [32] G. Artigas, P. López-Senín, C. González, N. Escaja, V. Marchán, *Org. Biomol. Chem.* **2015**, 13, 452–464.
- [33] J. D. McAnany, J. P. Reichert, B. L. Miller, *Bioorganic Med. Chem.* **2016**, 24, 3940–3946.
- [34] Z. Yang, Z. Fang, W. He, Z. Wang, H. Gan, Q. Tian, K. Guo, *Bioorg. Med. Chem. Lett.* **2016**, 26, 1671–1674.

- [35] M. Mondal, N. Radeva, H. Fanlo-Virgós, S. Otto, G. Klebe, A. K. H. Hirsch, *Angew. Chemie Int. Ed.* **2016**, *55*, 9422–9426.
- [36] P. Frei, L. Pang, M. Silbermann, D. Eris, T. Mühlethaler, O. Schwardt, B. Ernst, *Chem. Eur. J.* **2017**, *23*, 11570–11577.
- [37] J. Fu, H. Fu, M. Dieu, I. Halloum, L. Kremer, Y. Xia, W. Pan, S. P. Vincent, *Chem. Commun.* **2017**, *53*, 10632–10635.
- [38] J. Soubhye, M. Gelbcke, P. Van Antwerpen, F. Dufrasne, M. Y. Boufadi, J. Nève, P. G. Furtmüller, C. Obinger, K. Zouaoui Boudjeltia, F. Meyer, *ACS Med. Chem. Lett.* **2017**, *8*, 206–210.
- [39] A. G. Ekström, J. T. Wang, J. Bella, D. J. Campopiano, *Org. Biomol. Chem.* **2018**, *16*, 8144–8149.
- [40] M. Das, T. Yang, J. Dong, F. Prasetya, Y. Xie, K. H. Q. Wong, A. Cheong, E. C. Y. Woon, *Chem. – An Asian J.* **2018**, *13*, 2854–2867.
- [41] P. García, V. L. Alonso, E. Serra, A. M. Escalante, R. L. E. Furlan, *ACS Med. Chem. Lett.* **2018**, *9*, 1002–1006.
- [42] A. J. Clipson, V. T. Bhat, I. McNae, A. M. Caniard, D. J. Campopiano, M. F. Greaney, *Chem. – A Eur. J.* **2012**, *18*, 10562–10570.
- [43] M. Sindelar, K. T. Wanner, *ChemMedChem* **2012**, *7*, 1678–1690.
- [44] R. Nguyen, I. Huc, *Chem. Commun.* **2003**, 942–943.
- [45] D. A. Erlanson, A. C. Braisted, D. R. Raphael, M. Randal, R. M. Stroud, E. M. Gordon, J. A. Wells, *Proc. Natl. Acad. Sci.* **2000**, *97*, 9367–9372.
- [46] R. F. Ludlow, S. Otto, *J. Am. Chem. Soc.* **2008**, *130*, 12218–12219.
- [47] B. Shi, M. F. Greaney, *Chem. Commun.* **2005**, 886–888.
- [48] I. K. H. Leung, T. Brown, C. J. Schofield, T. D. W. Claridge, *Medchemcomm* **2011**, *2*, 390–395.
- [49] J. Newman, *Acta Crystallogr. Sect. D* **2004**, *60*, 610–612.
- [50] L. Reinhard, H. Mayerhofer, A. Geerlof, J. Mueller-Dieckmann, M. S. Weiss, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2013**, *69*, 209–214.
- [51] F. H. Niesen, H. Berglund, M. Vedadi, *Nat. Protoc.* **2007**, *2*, 2212–2221.
- [52] G. S. Sittampalam, N. P. Coussens, K. Brimacombe, A. Grossman, M. Arkin, D. Auld, C. Austin, J. Baell, B. Bejcek, J. M. M. Caaveiro, et al., **2004**.



- [53] T. Arakawa, Y. Kita, S. N. Timasheff, *Biophys. Chem.* **2007**, *131*, 62–70.
- [54] D. H. Rammner, *Ann. N. Y. Acad. Sci.* **1967**, *141*, 291–299.
- [55] M. Jackson, H. H. Mantsch, *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1991**, *1078*, 231–235.
- [56] A. Viegas, J. Manso, F. L. Nobrega, E. J. Cabrita, *J. Chem. Educ.* **2011**, *88*, 990–994.
- [57] B. Danieli, A. Giardini, G. Lesma, D. Passarella, B. Peretto, A. Sacchetti, A. Silvani, G. Pratesi, F. Zunino, *J. Org. Chem.* **2006**, *71*, 2848–2853.
- [58] C. Qiu, Z. Fang, L. Zhao, W. He, Z. Yang, C. Liu, K. Guo, *React. Chem. Eng.* **2019**, *4*, 658–662.
- [59] M. Jaegle, E. L. Wong, C. Tauber, E. Nawrotsky, C. Arkona, J. Rademann, *Angew. Chem. internat. Ed.* **2017**, *56*, 7358–7378.
- [60] A. M. Hartman, M. Mondal, N. Radeva, G. Klebe, A. K. Hirsch, *Int. J. Mol. Sci.* **2015**, *16*, DOI 10.3390/ijms160819184.
- [61] D. A. Erlanson, W. Jahnke, *Fragment-Based Drug Discovery: Lessons and Outlook*, Wiley-VCH Verlag GmbH, **2016**.
- [62] C. W. Murray, M. L. Verdonk, *J. Comput. Aided. Mol. Des.* **2002**, *16*, 741–753.